Achieving Secrecy without any Instantaneous CSI: Polar Coding for Fading Wiretap Channels

Hongbo Si
The University of Texas at Austin
HongboSi@utexas.edu

O. Ozan Koyluoglu
The University of Arizona
ozan@email.arizona.edu

Sriram Vishwanath
The University of Texas at Austin
sriram@austin.utexas.edu

***Abstract*—This paper presents a polar coding scheme for fading wiretap channels that achieves reliability as well as security without the knowledge of instantaneous channel state information at the transmitter. Specifically, a block fading model is considered for the wiretap channel that consists of a transmitter, a receiver, and an eavesdropper; and only the information regarding the statistics (i.e., distribution) of the channel state information is assumed at the transmitter. For this model, a coding scheme that hierarchically utilizes polar codes is presented in order to address channel state variation. In particular, on polarization of different binary symmetric channels over different fading blocks, each channel use (corresponding to a possibly different polarization) is modeled as an appropriate binary erasure channel over fading blocks. Polar codes are constructed for both coding over channel uses for each fading block and coding over fading blocks for certain channel uses. In order to guarantee security, message bits are transmitted such that they can be reliably decoded at the receiver, and random bits are introduced to exhaust the observations of the eavesdropper. It is shown that this coding scheme, without instantaneous channel state information at the transmitter, is secrecy capacity achieving for the corresponding fading binary symmetric wiretap channel.**

I. Introduction

Wiretap channels, introduced in the seminal paper of Wyner [1], model the communication between a transmitter and a receiver in the presence of an eavesdropper that overhears the transmitted signals via the channel between transmitter and eavesdropper. The transmitter's task is to hide information from the eavesdropper while communicating reliably to the receiver. Wyner studied this problem and characterized the capacity region when the eavesdropper is degraded [1]. The achievability scheme of [1] - adopted by numerous works in the literature - utilizes additional randomness in a random coding argument to eliminate leakage to the eavesdropper in order to achieve security. Yet, the design of secrecy achieving coding schemes with practical constraints such as low complexity and availability of channel state information remain as an important direction in the physical layer security.

Recently, polar codes have been utilized to enable capacity-achieving communication over degraded wiretap channels [2]–[5]. Polar codes are the first family of provably capacity achieving codes for symmetric binary-input discrete memoryless channels with low encoding and decoding complexity [6]. These codes rely on the "channel polarization" technique, which reconstructs a set of equivalent channels such that each of them is either purely noisy or noiseless. Noting that the fraction of noiseless channels approaches to the symmetric channel capacity, transmitting information symbols on the good instances and freezing the bad ones achieves the optimal rate. The schemes proposed in [2]–[5] are based on the behavior of polarization of degraded channels, where the polarized channels that are 'good' for the receiver and 'bad' for the eavesdropper are exploited to achieve secrecy. This scheme however requires instantaneous channel state information (CSI) at the transmitter, an assumption which may not hold in practice. In this work, we focus on relaxing this assumption, and develop polar coding schemes when only the statistics of CSI is known at the transmitter.

Recent studies on the design of polar coding schemes to achieve secrecy include [7]–[13], where strong security is considered in [3], [8], [11], [13], key agreement/generation is studied in [3], [9], [10], and other channel models different than discrete memoryless wiretap channel are considered in [4], [12]–[14]. Our model is similar to the fading models considered in [3], [14] but differentiates from all these prior studies in that only a statistical CSI for *both* receiver and eavesdropper channels is assumed at the transmitter. Polar coding schemes for fading wiretap channels are first studied in [3], where the transmitter has the knowledge of instantaneous CSI for the receiver's channel and statistical CSI for the eavesdropper's channel. With this setup, a key agreement scheme is proposed based on utilizing polar codes for each fading block, where the communicated bits over fading blocks are then used in a privacy amplification step to construct secret keys. This technique when combined with invertible extractors allows for secure message transmission but with the requirement of receiver CSI at the transmitter [3]. Recent work [14] proposes a polar coding scheme that utilizes artificial noise and multiple transmit antennas under the same assumption for the fading channels. However, only a guarantee of secrecy rate with some probability is achieved. In contrast, in this paper, we consider a fading channel model where the transmitter does not need to know *any instantaneous CSI*, but only its distribution for *both* receiver and eavesdropper channels. The hierarchical polar coding scheme proposed in this paper, to the best of our knowledge, is the first provably secrecy capacity achieving coding scheme for fading (binary symmetric) wiretap channels. Considering that this type of binary channels model the AWGN channels with BPSK modulation and demodulation, our framework covers a wide application scenarios in practice.

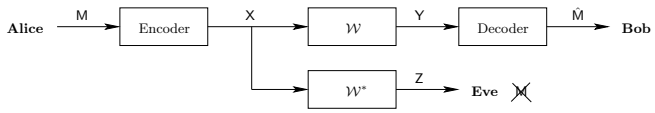

Fig. 1. System model for wiretap channels.

II. Problem Setup

Consider the fading (binary symmetric) wiretap channel model (Fig. 1): Alice desires to communicate with Bob through the main channel $\mathcal{W}$, and this transmission is also seen by an adversary (Eve) through the wiretap channel $\mathcal{W}^*$. Both these channels experience the following block fading phenomenon: With probability q_1, both channels are in a 'superior' state, where $\mathcal{W}$ behaves as BSC(p_1) and $\mathcal{W}^*$ behaves as BSC(p_1^*); and, with probability $q_2 \triangleq 1-q_1$, both channels are in a 'degraded' state, where $\mathcal{W}$ behaves as BSC(p_2) and $\mathcal{W}^*$ behaves as BSC(p_2^*). Hence, we have $p_1 \leq p_2 \leq 0.5$ and $p_1^* \leq p_2^* \leq 0.5$. Further, we assume that the wiretap channel is degraded compared to the main channel, i.e., $p_1 \leq p_1^*$, and $p_2 \leq p_2^*$.

In general, fading coefficients vary at a much slower pace than the transmission symbol duration. For such cases, block fading model is adopted, where the channel state is assumed to be constant within each coherence time interval, and follows a stationary ergodic process across fading blocks [15]. To this end, we consider a practically relevant scenario where channel state information (CSI) is available only at the decoder (CSI-D), while the encoder only knows the statistics of channel states. Under this model, a secrete message M is encoded by an encoding function $f(\cdot)$ to generate transmitted symbols: $\mathsf{X}_{1:NB} = f(\mathsf{M})$, where N is the length of a fading block, and B is the number of blocks. At the receiver, a decoding function $g(\cdot)$ gives an estimate of the estimate $\hat{\mathsf{M}}$, i.e., $\hat{\mathsf{M}} = g(\mathsf{Y}_{1:NB})$. The reliability of transmission is satisfied if

$$P_{\mathrm{e}} \triangleq \Pr\{\mathsf{M} \neq \hat{\mathsf{M}} | \mathsf{Y}_{1:NB}, \mathsf{S}\} \to 0, \text{ as } N, B \to \infty \quad (1)$$

where S denotes CSI, and (weak) security is defined as achieving

$$\frac{1}{NB} I(\mathsf{M}; \mathsf{Z}_{1:NB} | \mathsf{S}) \to 0, \text{ as } N, B \to \infty. \quad (2)$$

Under the degraded assumption, the secrecy capacity of the wiretap system can be upper bounded by

$$\begin{aligned} C^s_{\text{CSI-D}} &\triangleq \max_{p(x)} [I(\mathsf{X}; \mathsf{Y}|\mathsf{S}) - I(\mathsf{X}; \mathsf{Z}|\mathsf{S})] \\ &\overset{(a)}{\leq} \max_{p(x|s)} [I(\mathsf{X}; \mathsf{Y}|\mathsf{S}) - I(\mathsf{X}; \mathsf{Z}|\mathsf{S})] \\ &= \max_{p(x|1)} q_1 [I(\mathsf{X}; \mathsf{Y}|\mathsf{S}=1) - I(\mathsf{X}; \mathsf{Z}|\mathsf{S}=1)] \\ &\quad + \max_{p(x|2)} q_2 [I(\mathsf{X}; \mathsf{Y}|\mathsf{S}=2) - I(\mathsf{X}; \mathsf{Z}|\mathsf{S}=2)] \\ &\overset{(b)}{=} q_1 [H(p_1^*) - H(p_1)] + q_2 [H(p_2^*) - H(p_2)], \end{aligned} \quad (3)$$

where (a) follows by upper bounding the secrecy capacity with the case where encoder has CSI and adapts its coding scheme according to the channel states ($C^s_{\text{CSI-ED}}$) [16], and (b) is due to the secrecy capacity result for the degraded binary symmetric wiretap channel [17].

In this paper, assuming CSI is available only at the receivers, we provide a polar coding scheme that achieves this upper bound while satisfying reliability (1) and security (2) constraints. To this end, the upper bound (3) gives the secrecy capacity of our model. For the moment, we assume $p_1 \leq p_2 \leq p_1^* \leq p_2^*$, and the remaining case ($p_1 \leq p_1^* \leq p_2 \leq p_2^*$) can be treated similarly.

III. Polar Coding

The construction of polar codes is based on a phenomenon referred to as *channel polarization*. Consider a binary-input discrete memoryless channel $\mathcal{W}: \mathcal{X} \to \mathcal{Y}$, where $\mathcal{X} = \{0, 1\}$. Define $\boldsymbol{F} = [1, 0; 1, 1]$, and let $\boldsymbol{B}_N$ be the bit-reversal operator as defined in [6], where $N = 2^n$. By applying the transform $\boldsymbol{G}_N = \boldsymbol{B}_N \boldsymbol{F}^{\otimes n}$ ($\boldsymbol{F}^{\otimes n}$ denotes the n^{th} Kronecker product of $\boldsymbol{F}$) to $u_{1:N}$, the encoding is given by $x_{1:N} = u_{1:N}\boldsymbol{G}_N$, which is transmitted through N independent copies of $\mathcal{W}$. Now, consider N binary-input coordinate channels $\mathcal{W}_N^{(i)}: \mathcal{X} \to \mathcal{Y}^N \times \mathcal{X}^{i-1}$, where, for $i \in \{1, \ldots, N\}$, the transition probability is given by

$$\mathcal{W}_N^{(i)}(y_{1:N}, u_{1:i-1}|u_i) \triangleq \sum_{u_{i+1:N}} \frac{1}{2^{N-1}} \mathcal{W}^N(y_{1:N}|u_{1:N}\boldsymbol{G}_N).$$

Remarkably, as $N \to \infty$, the channels $\mathcal{W}_N^{(i)}$ polarize to either noiseless or pure-noisy, and the fraction of noiseless channels is close to $I(\mathcal{W})$, the symmetric capacity of channel $\mathcal{W}$ [6].

Given this polarization phenomenon, polar codes can be considered as $\boldsymbol{G}_N$-coset codes with parameters $(N, K, \mathcal{A}, u_{\mathcal{A}^c})$, where $u_{\mathcal{A}^c} \in \mathcal{X}^{N-K}$ is frozen vector (can be set to all-zeros for binary symmetric channels [6]), and the information set $\mathcal{A}$ is chosen as a K-element subset of $\{1, \ldots, N\}$ such that the Bhattacharyya parameters satisfy $Z(\mathcal{W}_N^{(i)}) \leq Z(\mathcal{W}_N^{(j)})$ for all $i \in \mathcal{A}$ and $j \in \mathcal{A}^c$, i.e., $\mathcal{A}$ denotes *good* channels. We use permutations (namely, ψ and ϕ in the sequel) to denote the increasing order of Bhattacharyya parameter values for the polarization of underlying channels.

A decoder for a polar code is the successive cancelation (SC) decoder, which gives an estimate $\hat{u}_{1:N}$ of $u_{1:N}$ given knowledge of $\mathcal{A}$, $u_{\mathcal{A}^c}$, and $y_{1:N}$ by computing

$$\hat{u}_i \triangleq \begin{cases} 1, & \text{if } i \in \mathcal{A}, \text{ and } \frac{\mathcal{W}_N^{(i)}(y_{1:N}, \hat{u}_{1:i-1}|1)}{\mathcal{W}_N^{(i)}(y_{1:N}, \hat{u}_{1:i-1}|0)} \geq 1, \\ 0, & \text{otherwise,} \end{cases}$$

in the order i from 1 to N. It has been shown that, by adopting an SC decoder, polar codes achieve any rate $R < I(\mathcal{W})$ with a decoding error scaling as $O(2^{-N^\beta})$, where $\beta < 1/2$. Moreover, the complexity for both encoding and decoding is $O(N \log N)$.

IV. A Hierarchical Polar Coding Scheme

A. Intuition

The intuition of hierarchical polar coding scheme originates from the polarization of degraded channels [18]. When polarizing two binary symmetric channels, the *good* channels

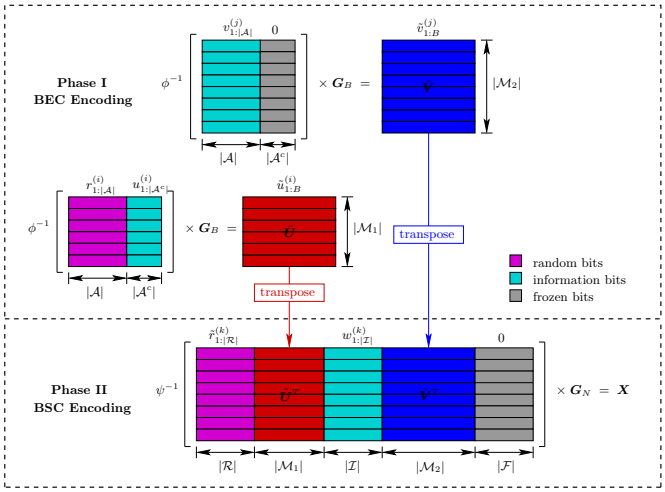

Fig. 2. Encoder of the polar coding scheme for wiretap channels.

of the polarized degraded channel is a subset of that of the superior one. [19] utilizes this property to construct hierarchical polar codes in order to achieve the capacity of fading binary symmetric channels. More precisely, polar codes are not only designed over channel uses within each fading block, but also utilized over different fading blocks. Inspired by this design, and combined with the polar coding scheme for wiretap channels [2]–[5], we design the proposed polar coding scheme for fading binary symmetric wiretap channels.

B. Encoder

The encoder works in two phases (see Fig. 2), hierarchically using polar codes to generate an NB-length codeword.

1) Phase I: BEC Encoding: Here, we consider two sets of messages to be encoded using polar encoders designed for binary erasure channels (BECs). For the first set of messages, we generate $|\mathcal{M}_1|$ number of BEC polar codes, where

$$|\mathcal{M}_1| = N[H(p_2^*) - H(p_1^*)]. \tag{4}$$

Consider a set of blockwise messages $u_{1:|\mathcal{A}^c|}^{(i)}$ with $i \in \{1, \ldots, |\mathcal{M}_1|\}$, where $\mathcal{A}$ is the information set for BEC(q_2), i.e.,

$$|\mathcal{A}| = B \cdot [q_1 - \epsilon], \tag{5}$$
$$|\mathcal{A}^c| = B \cdot [q_2 + \epsilon], \tag{6}$$

where ϵ is a positive number tending to 0 as N and B tend to infinity. For every $u_{1:|\mathcal{A}^c|}^{(i)}$, we combine it with $|\mathcal{A}|$ random bits to construct polar codeword $\tilde{u}_{1:B}^{(i)}$. Denoting the permutation for BEC(q_2) channel as ϕ, and the uniform random string as $r_{1:|\mathcal{A}|}^{(i)}$ (each bit is Ber$(1/2)$ distributed), the encoding process is given by

$$\tilde{u}_{1:B}^{(i)} = \boldsymbol{\mu}_{1:B}^{(i)} \times \boldsymbol{G}_B, \quad \phi\left(\boldsymbol{\mu}_{1:B}^{(i)}\right) = \left[\begin{array}{c|c} r_{1:|\mathcal{A}|}^{(i)} & u_{1:|\mathcal{A}^c|}^{(i)} \end{array} \right],$$

for every $i \in \{1, \ldots, |\mathcal{M}_1|\}$, where $\boldsymbol{G}_B$ is the polar generator matrix with size B. By collecting all $\tilde{u}_{1:B}^{(i)}$ together, the encoder generates a $|\mathcal{M}_1| \times B$ matrix $\tilde{\boldsymbol{U}}$. We denote $\tilde{\boldsymbol{U}}_k^T$ as the k-th row of the transpose of $\tilde{\boldsymbol{U}}$, where $k \in \{1, \ldots, B\}$.

Secondly, we generate $|\mathcal{M}_2|$ number of BEC polar codes, where

$$|\mathcal{M}_2| = N[H(p_2) - H(p_1)]. \tag{7}$$

Consider another set of blockwise messages $v_{1:|\mathcal{A}|}^{(j)}$ with $j \in \{1, \ldots, |\mathcal{M}_2|\}$. Each message is set as information bits to construct polar codeword $\tilde{v}_{1:B}^{(j)}$. More formally, this encoding process is given by

$$\tilde{v}_{1:B}^{(j)} = \boldsymbol{\nu}_{1:B}^{(j)} \times \boldsymbol{G}_B, \quad \phi\left(\boldsymbol{\nu}_{1:B}^{(j)}\right) = \left[\begin{array}{c|c} v_{1:|\mathcal{A}|}^{(j)} & 0 \end{array} \right],$$

for every $j \in \{1, \ldots, |\mathcal{M}_2|\}$. The collection of all $\tilde{v}_{1:B}^{(j)}$ together is denoted as a $|\mathcal{M}_2| \times B$ matrix $\tilde{\boldsymbol{V}}$. We denote $\tilde{\boldsymbol{V}}_k^T$ as the k-th row of the transpose of $\tilde{\boldsymbol{V}}$, where $k \in \{1, \ldots, B\}$.

2) Phase II: BSC Encoding: In this phase, we generate B number of BSC polar codes, each with length N. The encoded codewords from previous phase are embedded as messages of this phase. We consider a set of messages $w_{1:|\mathcal{I}|}^{(k)}$ with $k \in \{1, \ldots, B\}$, where

$$|\mathcal{I}| = N[H(p_1^*) - H(p_2)]. \tag{8}$$

For every $w_{1:|\mathcal{I}|}^{(k)}$, we introduce random bits $\tilde{r}_{1:|\mathcal{R}|}^{(k)}$, where

$$|\mathcal{R}| = N[1 - H(p_2^*) - \epsilon], \tag{9}$$

and combine the output from the previous phase as message to construct polar codeword $x_{1:N}^{(k)}$. More formally, if we denote the reordering permutation for BSC(p_1) as ψ, then the encoder of this phase can be expressed as

$$x_{1:N}^{(k)} = \boldsymbol{\omega}_{1:N}^{(k)} \times \boldsymbol{G}_N,$$
$$\psi\left(\boldsymbol{\omega}_{1:N}^{(k)}\right) = \left[\begin{array}{c|c|c|c|c} \tilde{r}_{1:|\mathcal{R}|}^{(k)} & \tilde{\boldsymbol{U}}_k^T & w_{1:|\mathcal{I}|}^{(k)} & \tilde{\boldsymbol{V}}_k^T & 0 \end{array} \right],$$

for every $k \in \{1, \ldots, B\}$, where $\boldsymbol{G}_N$ is the polar generator matrix with size N. That is, the codewords generated from BEC encoding phase are transposed and embedded into the messages of the BSC encoding process. We denote these codewords by a $B \times N$ matrix $\boldsymbol{X}$. The proposed encoder is illustrated in Fig. 2.

C. Decoder for the Main Channel

The codewords $x_{1:N}^{(k)}$ are transmitted through both the main channel and the wiretap channel. After receiving the output sequence $y_{1:N}^{(k)}$ for all $k \in \{1, \ldots, B\}$, the task of the decoder is to make estimates for all the information and random bits. In particular, the decoder aims to recover $u_{1:|\mathcal{A}^c|}^{(i)}$, $v_{1:|\mathcal{A}|}^{(j)}$, $w_{1:|\mathcal{I}|}^{(k)}$, $r_{1:|\mathcal{A}|}^{(i)}$, and $\tilde{r}_{1:|\mathcal{R}|}^{(k)}$ successfully with high probability. As that of the encoding process, the decoding process also works in phases.

1) Phase I: BSC Decoding for the Superior Channel State: In this phase, using the BSC SC decoder, channels corresponding to the superior state are decoded. More precisely, since the receiver knows the channel states, it can adopt the correct SC decoder to obtain estimates $\hat{\boldsymbol{\omega}}_{1:N}^{(k)}$ from $y_{1:N}^{(k)}$ for every k corresponding to the superior channel state. To this end, the

decoder adopted in this phase is the classical BSC SC polar decoder with parameter p_1, i.e.,

$$\hat{\boldsymbol{\omega}}_n^{(k)} = \begin{cases} 1, & \text{if } n \notin \mathcal{F},\ \frac{\mathcal{W}_{1,N}^{(n)}(y_{1:N}^{(k)}, \hat{\boldsymbol{\omega}}_{1:n-1}^{(k)}|1)}{\mathcal{W}_{1,N}^{(n)}(y_{1:N}^{(k)}, \hat{\boldsymbol{\omega}}_{1:n-1}^{(k)}|0)} \geq 1, \\ 0, & \text{otherwise,} \end{cases}$$

in the order n from 1 to N, and $\mathcal{W}_{1,N}^{(n)}$ is the n-th polarized channel from BSC(p_1). Then, for every k corresponding to the superior channel state, the decoder can obtain the messages (with the knowledge of the frozen symbols corresponding to $\mathcal{F}$ indices)

$$\psi\left(\hat{\boldsymbol{\omega}}_{1:N}^{(k)}\right) = \left[\begin{array}{ccccccc} \hat{\tilde{r}}_{1:|\mathcal{R}|}^{(k)} & | & \hat{\tilde{\boldsymbol{U}}}_k^T & | & \hat{w}_{1:|\mathcal{I}|}^{(k)} & | & \hat{\tilde{\boldsymbol{V}}}_k^T \quad | \quad 0 \end{array}\right].$$

However, for the blocks with degraded channel states, one cannot decode reliably because the frozen bits corresponding to set $\mathcal{M}_2$ are unknown at the decoder. At this point, we use the next phase to decode these frozen bits using a BEC SC decoder. To proceed, we construct a $B \times |\mathcal{M}_2|$ matrix $\hat{\tilde{\boldsymbol{V}}}^T$ such that its rows corresponding to the superior state are determined in previous decoding process, while the ones corresponding to the degraded states are all set to erasures.

2) Phase II: BEC Decoding: In this phase, we decode the frozen bits with respect to the degraded channel state. More precisely, each row of matrix $\hat{\tilde{\boldsymbol{V}}}$, denoted by $\hat{\tilde{\boldsymbol{V}}}_j$ for $j \in \{1, \ldots, |\mathcal{M}_2|\}$, is considered as the input to the decoder, and the receiver aims to obtain an estimate of the information bits from it using BEC SC decoder. To this end, the decoder adopted in this phase is the classical BEC SC decoder with parameter q_2, i.e.,

$$\hat{\boldsymbol{\nu}}_b^{(j)} = \begin{cases} 1, & \text{if } b \in \mathcal{A}, \text{ and } \frac{\mathcal{W}_{\mathrm{e},B}^{(b)}(\hat{\tilde{\boldsymbol{V}}}_j, \hat{\boldsymbol{\nu}}_{1:b-1}^{(j)}|1)}{\mathcal{W}_{\mathrm{e},B}^{(b)}(\hat{\tilde{\boldsymbol{V}}}_j, \hat{\boldsymbol{\nu}}_{1:b-1}^{(j)}|1)} \geq 1, \\ 0, & \text{otherwise,} \end{cases}$$

in the order b from 1 to B, and $\mathcal{W}_{\mathrm{e},B}^{(b)}$ is the b-th polarized channel from BEC(q_2). Then, for every b, the decoder can declare $\phi\left(\hat{\boldsymbol{\nu}}_{1:B}^{(j)}\right) = \left[\begin{array}{ccc} \hat{v}_{1:|\mathcal{A}|}^{(j)} & | & 0 \end{array}\right]$. At this point, the decoder can reconstruct all erased bits as well. More precisely, the erased rows in $\hat{\tilde{\boldsymbol{V}}}^T$ can be recovered, and they can be further utilized to decode the information bits in blocks with the degraded channel state in the next phase.

3) Phase III: BSC Decoding for the Degraded Channel State: In this phase, the remaining blocks from Phase I are decoded by using BSC SC decoders with respect to degraded channel states. In particular, bits in the frozen set for the degraded channel state (set $\mathcal{F}$ and set $\mathcal{M}_2$) are known due to the previous phases. Hence, the receiver can decode from $y_{1:N}^{(k)}$ using BSC SC decoder with parameter p_2, i.e.,

$$\hat{\boldsymbol{\omega}}_n^{(k)} = \begin{cases} 1, & \text{if } n \notin \mathcal{F} \cup \mathcal{M}_2,\ \frac{\mathcal{W}_{2,N}^{(n)}(y_{1:N}^{(k)}, \hat{\boldsymbol{\omega}}_{1:n-1}^{(k)}|1)}{\mathcal{W}_{2,N}^{(n)}(y_{1:N}^{(k)}, \hat{\boldsymbol{\omega}}_{1:n-1}^{(k)}|0)} \geq 1, \\ \hat{\tilde{\boldsymbol{V}}}_{kn}^T, & \text{if } n \in \mathcal{M}_2, \\ 0, & \text{otherwise,} \end{cases}$$

in the order n from 1 to N, and $\mathcal{W}_{2,N}^{(n)}$ is the n-th polarized channel from BSC(p_2). Then, for every k corresponding to the degraded channel state, the decoder declares

$$\psi\left(\hat{\boldsymbol{\omega}}_{1:N}^{(k)}\right) = \left[\begin{array}{ccccccc} \hat{\tilde{r}}_{1:|\mathcal{R}|}^{(k)} & | & \hat{\tilde{\boldsymbol{U}}}_k^T & | & \hat{w}_{1:|\mathcal{I}|}^{(k)} & | & \hat{\tilde{\boldsymbol{V}}}_k^T \quad | \quad 0 \end{array}\right].$$

Hence, after this decoding procedure, the receiver makes an estimate $\hat{\tilde{\boldsymbol{U}}}$ of matrix $\tilde{\boldsymbol{U}}$, which further implies all information bits in $u_{1:|\mathcal{A}^c|}^{(i)}$ are decoded. Note that, in addition to information bits, all random bits are decoded reliably as well. However, in order to guarantee security, we set these bits random (instead of information).

D. Achievable Rate and Reliability

The proposed hierarchical scheme allows for recovering all information bits (represented by light blue in Fig. 2) reliably, as long as the designed rates of polar codes do not exceed the corresponding channel capacities. Hence, the achievable rate is given by

$$\begin{aligned} R &= \frac{1}{NB}\left(|\mathcal{M}_2| \times |\mathcal{A}| + |\mathcal{M}_1| \times |\mathcal{A}^c| + B \times |\mathcal{I}|\right) \\ &= [H(p_2) - H(p_1)] \times [q_1 - \epsilon] \\ &\quad + [H(p_2^*) - H(p_1^*)] \times [q_2 + \epsilon] + [H(p_1^*) - H(p_2)] \\ &= [H(p_1^*) - H(p_1)] \times q_1 + [H(p_2^*) - H(p_2)] \times q_2 - \delta(\epsilon), \end{aligned}$$

where we have used (4), (5), (6), (7), and (8), and $\delta(\epsilon) \to 0$ as N, $B \to \infty$. In this scheme, B number of N-length polar codes are decoded in Phase I and III in total, and $|\mathcal{M}_2|$ number of B-length polar codes are decoded in Phase II. Hence, the decoding error probability is upper bounded by

$$\Pr\{\mathsf{M} \neq \hat{\mathsf{M}} | \mathsf{Y}_{1:NB}, \mathsf{S}\} \leq B \cdot 2^{-N^\beta} + |\mathcal{M}_2| \cdot 2^{-B^\beta}, \quad (10)$$

where $\beta < 1/2$; and, M is the collection of random variables representing for all information bits (its realizations include $u_{1:|\mathcal{A}^c|}^{(i)}$, $v_{1:|\mathcal{A}|}^{(j)}$, and $w_{1:|\mathcal{I}|}^{(k)}$), and $\hat{\mathsf{M}}$ is the estimate of M obtained at the legitimate receiver. Noting that the right hand side of (10) tends to 0 when implemented with properly large B and N, the proposed scheme achieves the upper bound given by (3) reliably.

E. Security

Assume that, in addition to $z_{1:N}^{(k)}$, a genie reveals Eve all information bits $u_{1:|\mathcal{A}^c|}^{(i)}$, $v_{1:|\mathcal{A}|}^{(j)}$, and $w_{1:|\mathcal{I}|}^{(k)}$. Under this condition, we show that all random bits can be reliably decoded at Eve. More precisely, the decoder designed for the eavesdropper also works in phases, similar to the one for the main channel. In Phase I, the decoder still works over the blocks with the superior channel state. However, for the wiretap channel with superior channel state, the frozen set consists of bits not only in set $\mathcal{F}$, but also in sets $\mathcal{M}_2$ and $\mathcal{I}$. Since we have assumed the information bits are known at Eve, the classical BSC(p_1^*) SC decoder can be used to decode the random bits. Then, in the next phase, we aim to recover the unknown frozen bits corresponding to the degraded channel state, where a similar scheme as that of the main receiver is adopted. More precisely, we utilize the BEC(q_2) SC decoder over each row of the matrix after transpose. This scheme successively recovers the erased elements, as the frozen bits for this BEC is the information

bits $u_{1:|\mathcal{A}^c|}^{(i)}$ and they are assumed to be known. Finally, the decoded result from the BEC decoding phase is utilized at the BSC decoding for the degraded state, where the classical BSC(p_2^*) SC decoder is adopted.

By adopting this hierarchical polar decoder, Eve can decode all random bits with high probability, i.e.,

$$\Pr\{\mathsf{R} \neq \hat{\mathsf{R}} | \mathsf{Z}_{1:NB}, \mathsf{M}, \mathsf{S}\} \leq B \cdot 2^{-N^\beta} + |\mathcal{M}_1| \cdot 2^{-B^\beta}, \quad (11)$$

where R is the collection of random variables representing for random bits (its realization include $r_{1:|\mathcal{A}|}^{(i)}$ and $\tilde{r}_{1:|\mathcal{R}|}^{(k)}$), and $\hat{\mathsf{R}}$ is the estimate of R. Then, using Fano's inequality, together with (11), we have

$$\begin{aligned} H(\mathsf{R}|\mathsf{Z}_{1:NB}, \mathsf{M}, \mathsf{S}) &\leq H(B \cdot 2^{-N^\beta} + |\mathcal{M}_1| \cdot 2^{-B^\beta}) \\ &+ [B \cdot 2^{-N^\beta} + |\mathcal{M}_1| \cdot 2^{-B^\beta}] \cdot [|\mathcal{R}| \cdot B + |\mathcal{A}| \cdot |\mathcal{M}_1|]. \quad (12) \end{aligned}$$

Based on this, the following steps provide an upper bound (omitting the subscript of Z):

$$\begin{aligned} I(\mathsf{M}; \mathsf{Z}|\mathsf{S}) &= I(\mathsf{M}, \mathsf{R}; \mathsf{Z}|\mathsf{S}) - [H(\mathsf{R}|\mathsf{M}, \mathsf{S}) - H(\mathsf{R}|\mathsf{Z}, \mathsf{M}, \mathsf{S})] \\ &\overset{(a)}{=} I(\mathsf{M}, \mathsf{R}; \mathsf{Z}|\mathsf{S}) - H(\mathsf{R}) + H(\mathsf{R}|\mathsf{Z}, \mathsf{M}, \mathsf{S}) \\ &\overset{(b)}{\leq} NB \cdot C_{\text{CSI-D}}(\mathcal{W}^*) - H(\mathsf{R}) + H(\mathsf{R}|\mathsf{Z}, \mathsf{M}, \mathsf{S}) \\ &\overset{(c)}{=} NB \cdot C_{\text{CSI-D}}(\mathcal{W}^*) - B \cdot |\mathcal{R}| - |\mathcal{A}| \cdot |\mathcal{M}_1| \\ &\quad + H(\mathsf{R}|\mathsf{Z}, \mathsf{M}, \mathsf{S}) \\ &\overset{(d)}{=} NB \cdot C_{\text{CSI-D}}(\mathcal{W}^*) - B \cdot N[1 - H(p_2^*) - \epsilon] \\ &\quad - B[q_1 - \epsilon] \cdot N[H(p_2^*) - H(p_1^*)] \\ &\quad + H(\mathsf{R}|\mathsf{Z}, \mathsf{M}, \mathsf{S}) \\ &\overset{(e)}{\leq} H(\mathsf{R}|\mathsf{Z}, \mathsf{M}, \mathsf{S}) - NB \cdot \delta'(\epsilon), \end{aligned}$$

where $\delta'(\epsilon) \to 0$ as N, $B \to \infty$, and (a) follows as R is independent of M and S; (b) is due to the definition of channel $\mathcal{W}^*$'s capacity with CSI-D; (c) is due to the assumption that R is uniform; (d) is due to equations (5), (4), and (9); (e) is due to the ergodic capacity of the fading eavesdropper channel with CSI known only at the decoder, i.e., $C_{\text{CSI-D}}(\mathcal{W}^*) \leq C_{\text{CSI-ED}}(\mathcal{W}^*) = q_1[1 - H(p_1^*)] + q_2[1 - H(p_2^*)]$.

Finally, combining this with the result of (12), we have $\frac{1}{NB} I(\mathsf{M}; \mathsf{Z}_{1:NB}|\mathsf{S}) \to 0$, as N and B tends to infinity (with proper choice of the their scaling relationship). Hence, the proposed scheme achieves the secrecy constraint.

V. Conclusion

In this paper, a hierarchical polar coding scheme is presented for the binary symmetric wiretap channel with block fading. By utilizing an erasure decoding approach at the receiver, this scheme uses the polarization of degraded binary symmetric channels to compensate for the impact of channel fading. Simultaneously, to combat eavesdropping, random bits are injected into the encoded symbols, and the resulting coding scheme is shown to achieve the secrecy capacity of this channel. The proposed coding scheme requires long codeword lengths to make the error probability arbitrarily small. This requirement translates to requiring long coherence intervals and large number of fading blocks as our approach utilizes coding over both channel uses and fading blocks. Therefore, our coding scheme fits to the fading channels with moderate/long coherence time and large number of fading blocks. Although we consider binary symmetric channels with only two fading states in this paper, the hierarchical polar coding scheme can be applied as a general method to other scenarios (such as fading blocks with more states) for simultaneously resolving fading and security problems. Its application to strong security will be reported as future work.

References

[1] A. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[2] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011.

[3] O. O. Koyluoglu and H. El Gamal, "Polar coding for secure transmission and key agreement," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1472–1483, Oct. 2012.

[4] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested polar codes for wiretap and relay channels," *IEEE Communications Letters*, vol. 14, no. 8, pp. 752–754, Aug. 2010.

[5] E. Hof and S. Shamai, "Secrecy-achieving polar-coding," in *Proc. 2010 IEEE Information Theory Workshop (ITW 2010)*, Aug. 2010, pp. 1–5.

[6] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[7] E. Abbe, "Low complexity constructions of secret keys using polar coding," in *Proc. 2012 IEEE Information Theory Workshop (ITW 2012)*, Sep. 2012, pp. 1–5.

[8] E. Sasoglu and A. Vardy, "A new polar coding scheme for strong security on wiretap channels," in *Proc. 2013 IEEE International Symposium on Information Theory Proceedings (ISIT 2013)*, Jul. 2013, pp. 1117–1121.

[9] D. Sutter, J. M. Renes, and R. Renner, "Efficient one-way secret-key agreement and private channel coding via polarization," *Advances in Cryptology - ASIACRYPT 2013*, vol. 8269, pp. 194–213, 2013.

[10] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," in *Proc. 2013 IEEE Information Theory Workshop (ITW 2013)*, Sep. 2013, pp. 1–5.

[11] Y.-P. Wei and S. Ulukus, "Polar coding for the general wiretap channel," *arXiv:1410.3812*, Oct. 2014.

[12] R. A. Chou and M. R. Bloch, "Polar coding for the broadcast channel with confidential messages and constrained randomization," *arXiv:1411.0281*, Nov. 2014.

[13] T. C. Gulcu and A. Barg, "Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component," *arXiv:1410.3422*, Oct. 2014.

[14] M. Zheng, M. Tao, and W. Chen, "Polar coding for secure transmission in MISO fading wiretap channels," *arXiv:1411.2463*, Nov. 2014.

[15] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[16] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[17] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge University Press, 2011.

[18] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, École Polytechnique Fédérale de Lausanne, 2009.

[19] H. Si, O. O. Koyluoglu, and S. Vishwanath, "Polar coding for fading channels: binary and exponential channel cases," *IEEE Transactions on Communications*, vol. 62, no. 8, pp. 2638–2650, Aug. 2013.